Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.18.(2) 2026,pp.Comp 50–61

Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Quality of Experience in Multimedia Streaming: Objective Metrics, Subjective Assessment, and Trends for Research and Practice

Ahmed M. Kareem[*,1] ,*Muntadher Khamees*[2], *Alaa Taima Albu-Slaih*[3]

[1]*Computer Science, University of Diyala, Iraq, Email: scicomphd222304@uodiyala.edu.iq*

[2]*Computer Science, University of Diyala, Iraq, Email: alkarawis@uodiyala.edu.iq*

[3]*Computer Science, University of Al-Qadisiyah, Iraq, Email: alaa.taima@qu.edu.iq*

ARTICLE INFO

Article history:
Received: 21 /01/2026
Rrevised form: 22 /02/2026
Accepted : 23 /02/2026
Available online: 30 /06/2026

Keywords:

Quality of Experience (QoE), Quality of Service (QoS), subjective assessment, Mean Opinion Score (MOS), Absolute Category Rating (ACR), HTTP adaptive streaming (DASH/HLS), VMAF, ITU-T P.1203/P.1204

ABSTRACT

Quality of Experience (QoE) has become the dominant success criterion for modern multimedia streaming, yet the literature is fragmented across network QoS studies, perceptual quality models, subjective user experiments, and data-driven prediction. This survey consolidates these strands and explains how to translate QoE concepts into measurable, comparable, and actionable engineering signals. Specifically, we (i) propose a taxonomy linking network-, application-, and perceptual-level metrics to user-perceived outcomes; (ii) synthesize evidence on how objective indicators relate to subjective ratings (MOS/ACR) in adaptive streaming scenarios (startup delay, rebuffering, and quality switching); (iii) compare widely used models and standards (e.g., ITU-T P.1203/P.1204 and VMAF) with their assumptions and limitations; and (iv) summarize datasets, protocols, and practical guidance for reproducible evaluation and machine learning-based QoE prediction. Finally, we highlight open gaps—live-streaming QoE, cross-device perception, and model generalization under distribution shift—and outline research directions for 5G/6G, edge computing, and immersive media.

MSC..

https://doi.org/10.29304/jqcsm.2026.18.22601

Introduction

Quality of Experience (QoE) denotes how end users perceive and evaluate multimedia streaming, and it has become a key determinant of platform success, influencing retention, engagement, and revenue. In contrast, Quality of Service (QoS) describes technical delivery conditions, including throughput, latency, jitter, packet loss, and rebuffering; QoE is fundamentally user-centric, dependent on human perception, content characteristics, device limitations, and the usage context. Consequently, identical QoS values can yield varying QoE across devices (mobile versus television), applications (Video on Demand versus live streaming), and viewing contexts (commuting versus home).

∗Corresponding author: *Ahmed M. Kareem*

Email addresses: *scicomphd222304@uodiyala.edu.iq*

Communicated by 'sub etitor'

The research landscape has broadened rapidly, with standardized parametric models (for example, ITU-T P.1203/P.1204), perceptual metrics such as VMAF, large-scale field telemetry, and machine learning–based QoE prediction widely employed, while subjective studies remain essential for validation. However, many existing surveys tend to emphasize either network management, subjective methodology, or specific QoE models, leaving practitioners without an integrated perspective on (i) what to measure, (ii) how to relate objective indicators to perception, and (iii) how to evaluate and reproduce results across datasets and platforms. This paper addresses this gap by synthesizing objective metrics, subjective assessment methods, and emerging trends, with an emphasis on practical guidance for research and practice.

1.1 Novelty and Contributions of This Survey

- Positioning relative to prior QoE surveys: a direct comparison of major surveys and clarification of what is novel in scope and synthesis.

- Unified taxonomy: a mapping of QoE indicators across layers (network, application, perceptual) and their connection to user-perceived outcomes.

- Critical synthesis: an evaluation of the strengths and limitations of metrics and models (e.g., ITU-T P.1203/P.1204, VMAF), highlighting scenarios where they fail (live streaming, short clips, mobile viewing).

- Practical guidance: a concise presentation of datasets, benchmarking protocols, and reporting checklists to enable reproducible QoE evaluation and machine learning–based modeling.

- Research gaps and directions: identification of open problems such as distribution shift, cross-device calibration, and immersive/edge scenarios. See Table 1 presents representative QoE surveys and clarifies the positioning of this work.

Table 1: QoE Surveys

Survey / Reference	Year	Primary scope	Covers objective metrics	Covers subjective methods	Practical guidance / datasets
Alreshoodi & Woods, "Survey on QoE–QoS correlation models"	2013	QoE–QoS mapping models	✓	✓	Limited
Barakabitze et al., "QoE management of multimedia streaming services…"	2020	QoE management (monitoring/optimization)	✓	Partial	✓
Kougioumtzidis et al., survey on QoE assessment & ML prediction	2022	QoE assessment + ML prediction	✓	Partial	Partial
Zhou et al., "A brief survey on adaptive video	2022	HAS QoE models	✓	✓ (databases)	Partial

streaming QoE assessment"					
This survey	2026	Integrated metrics + perception + practice	✓ (taxonomy + comparison)	✓ (standards + design)	✓ (tables + checklists)

2. Conceptual Foundations of Quality of Experience

Quality of experience (QoE) refers to the quality perceived by the user, not the technical quality of multimedia content or the quality of service (QoS) from the network that supports it. Thus, QoE emphasizes a user-centric definition of quality rather than an all-encompassing application, network, or technology monitoring. Therefore, QoE extends some traditional definitions of quality. A complete survey on user-perceived quality in multimedia streaming would separate video and audio quality perceptions, physical and technical quality, semantic quality, and other types of perceived quality. Traditional QoS metrics are network-related quantities such as throughput, latency, jitter, and packet loss; they are different technological parameters for networked services. These metrics do not depend directly on the user but on the amount of information that is networked (Kougioumtzidis et al., 2022; Barakabitze et al., 2020).

Quality perceived by users has two very closely related but distinct components: perceptual quality and temporal quality. Perceptual quality refers to the intrinsic value of content delivered or streamed; further elaboration on concepts like quality-of-perception and perceptual-relevance will enhance this analysis .

3. Objective Metrics for QoE in Streaming

A wide variety of objective metrics are available for evaluating the Quality of Experience (QoE) provided by multimedia streaming services. These can be classified into network-level, application-level, and user-perceivable quality indicators. Network-level metrics relate to the transport layer and hence provide an indirect view regarding what Quality of Service (QoS) is being experienced by end users. Application-level metrics describe how well the streaming application behaves with respect to the network and how much it fulfills user expectations regarding service provision. Quality indicators are often correlated with both QoS as well as user-perceived quality; they may be derived from encoding and decoding operations, characteristics related to video content itself, or through interaction with users. A single objective measure may suffice for some applications while more sophisticated assessment becomes necessary in other contexts. The availability of several complementary measurements along with relevant metrics enhances QoE monitoring as well scientific investigations plus machine learning applications.

Network-level metrics comprise parameters related to the operation of the Transmission Control Protocol (TCP) and Quick UDP Internet Connections (QUIC), such as round-trip time and congestion window. These metrics have an impact on QoE through parameters like throughput, latency, jitter, and packet loss. A minimum standard of QoS is necessary but placing too much importance on network-related measurements may take away from the end-user perspective. Applications are now adaptive and continuously change according to many conditions; thus focusing on other parameters more closely related to service delivery could improve this focus. Application-level metrics such as bandwidth, buffer, and playout duration stress the importance of monitoring at the application layer. In video streaming, quality in terms of visual experience directly translates to overall experience making perceptual video quality models relevant. For audio streaming, there are usually only a few discrete bitrates that define the service so modeling experience using audio quality indicators at the decoder may be more relevant.

The overview of common objective QoE indicators, indicating where they are measured and typical limitations, see Table 2. Key caution: objective indicators are most dependable when accompanied by contextual information (content type, device class, and streaming mode such as VoD or live). For instance, throughput may appear sufficient, yet short-term fluctuations can lead to buffer depletion and stalls; likewise, frame-level fidelity metrics can be high even when QoE is poor due to repeated quality switches. Consequently, best practice entails combining at least one temporal-interruption metric (startup/stall), one adaptation-stability metric (switch rate/amplitude), and one perceptual-quality metric (e.g., VMAF or an ITU model), and validating the correspondence to subjective scores on a representative dataset.

Table 2: Objective QoE Indicators

Category	Example indicators	Measurement layer	What it captures well	Common limitations / caveats
Network (QoS)	RTT, jitter, packet loss, throughput/goodput	Transport / network	Capacity & congestion signals	Indirect; weakly maps to perception without context
Application / session	startup delay, rebuffer count & duration, buffer level, bitrate switches, stall ratio	Player / client	Directly affects interruptions and responsiveness	Implementation-dependent; logging definitions vary
Content / perceptual	PSNR, SSIM, VMAF; bitrate/resolution/frame-rate	Decoder / media	Visual fidelity of decoded frames	Full-reference metrics need source; may miss temporal artifacts
Hybrid standardized models	ITU-T P.1203 / P.1204 outputs	Client-side telemetry + metadata	Integrates stalling + representation quality into a single QoE score	Assumptions about codec/device; calibration needed for new settings
Learned predictors	ML/DL QoE regressors using QoS+player logs	Data-driven	Can fit complex nonlinear perception patterns	Generalization risk; requires representative datasets and careful validation

3.1. Network-Level Metrics

Quality of Experience determines the level of streaming satisfaction perceived from multimedia content consumed in a streaming way which depends on the performance of some components of delivery systems. There have been many works about assessing QoE in multimedia streaming. Since objective metrics can provide fast feedback to engineers during design stages and online operations, they are increasingly being demanded for analyzing and monitoring content provided through streaming platforms . However, subjective assessment will still be necessary in order to relate these metrics with quality perceived by end users and validate that simulated environments give realistic responses. The concepts behind QoE will be explained in subsequent sections before presenting objective QoE metrics in multimedia streaming.

3.2 Application-Level Metrics

The quality of multimedia streaming as perceived by users depends on system/application behavior as well as end-user perspective. The former refers to configuration decisions made by engineers or algorithm designers—video/audio encoding tasks or adaptation logic—affecting application-layer streaming metrics that impact Quality of Experience . Five categories are relevant here: those related to bitrate-ladders, starting-up lag, stalling, and adaptation. Network- and streaming-application-layer QoE parameter analysis is important for both academia and industry.

Bitrate-ladder-related parameters are systematic choices about the audio/video codec and the bitrate ladder at the transmission side. The choice of codecs has a direct impact on content semantic quality, and therefore, the codec–bitrate correlation should be analyzed in-depth. The full usage of the entire bitrate ladder indicates less consideration of the QoE at user-side terminals. Implementing a complete ladder at the user terminal is only possible through HLS and DASH mechanisms. The adaptation procedure using segments and changes in the profile used during playback can significantly impact quality perceived by users still falls under research in content-adaptive streaming solutions.

3.3. Perceptual and Quality Metrics

Perceptual and quality metrics are essential for video and image quality assessment Torres Vega et al., 2018. Various methodologies are available including no-reference metrics, structure distortion measurement, and hybrid metrics. Peak Signal-to-Noise Ratio (PSNR) is a standardized method along with some recent metrics that consider structural as well as perceptual distortions. Objective assessments are compared with subjective user experience to ensure reliability; therefore, benchmarked against subjective user experience for reliability in Perceptual and quality metrics applied to streaming services HTTP adaptive streaming multimedia applications help improve Quality of Experience.

4. Subjective Assessment

Subjective Quality of Experience (QoE) assessment elucidates how users truly perceive impairments such as stalls, quality switches, and latency, supplying the ground truth necessary to validate objective metrics and learned predictors. Given that QoE depends on human perception and context, subjective studies remain indispensable—particularly when new codecs, devices, or streaming scenarios are introduced.

4.1. Controlled Experiments and Study Design

Designing controlled experiments for Quality of Experience (QoE) has challenges in participant, stimulus, and methodology selection. Constraints from time, budget, or other resources are common among researchers. However, the basic principles of experimental design should be followed to ensure that results are meaningful. Industrial realism and external validity criteria take precedence over aesthetic considerations (Venkataraman Iyengar, 2011). To ensure the reproducibility of results, protocols for participant recruitment, experimental stimulation, and setting have been defined. These guidelines are not exhaustive but serve as a starting point for designing interesting and impactful QoE activities. The objectives of any experiment are relevant data acquisition with broad interest and engagement resulting in statistically reliable outcomes; thus shaping choices regarding demography of participants, source of stimulation, and apparatus used. After stating its objectives, an experiment selects candidate materials and signals applicable to it and compliant with conditions laid down previously. In the absence of fixed-duration source materials, accurate knowledge about total duration is required for calculating temporal variables; calculated durations should simulate system-level playback plus any other timing elements such as silence periods for audio that do not contribute to overall experience yet add up to total time. Stop, pause or seek commands engage different internal apparatus subprocesses whose timing may add significantly to delay—and hence experience.

The summary of commonly employed subjective methodologies and the corresponding standards that define them, See Table 3. For streaming systems, subjective design should explicitly account for typical events such as startup delay, rebuffering, and quality switching, and document viewing conditions (display size, viewing distance, audio). When the objective–subjective correlation is the objective, the study should log precise player events and segment qualities to enable reproducible model training and comparison.

Table 3: Subjective Methodologies

Method / scale	Typical use	Standard guidance	Strengths	Caveats
MOS (1–5) / ACR	Overall quality ratings per	ITU-T P.910	Simple, widely reported; good	Sensitive to context; needs enough

	clip/segment		comparability	participants
DSIS / DSCQS	Reference vs impaired comparison	ITU-R BT.500	Reduces bias using reference; strong for codec comparisons	Less realistic for adaptive streaming with stalls/switches
SAMVIQ / continuous scales	Fine-grained ratings with replay	ITU-R BT.500 (options)	Higher resolution judgments	Longer experiments; fatigue risk
Crowdsourced QoE	Large-scale, diverse users/devices	Guidance varies; often adapted from ITU methods	Scalable and cost-effective	Harder control of viewing conditions; requires screening

4.2. Rating Scales and Data Analysis

Keeping the order of ratings across subjects preserves some biases particularly when the subjects watch the same sequences or items; this risk is avoided if an experiment is balanced in which case each item has a unique exclusive pairing with every other item in the set provided that there are more than two items to be rated. The presentation order within such a design uses either random arrangements or is counter-balanced explicitly.

Mean Opinion Score and Absolute Category Rating are the two most common methods of subjective assessment used in multimedia streaming. The rating scale for MOS is a single score between 1 and 5, with 1 being the worst quality and 5 being the best quality. It can be administered in groups or individually as single-choice responses. Since MOS is widely used to represent implicit preference levels in different video analysis sections, we will use it throughout this study. ACR allows for a rating scale between 0 and 10 while the sequence remains fixed; thus, denoting all-out perceived quality rather than preference, it may produce either a floating or discrete result after several observations of the same sequence.

The data collected is usually analyzed statistically using Statistical Analysis System (SAS), Analysis of Variance (ANOVA), or any other embedded systems that may correspondingly be applied. To ensure reliability of results, Control-Check Questions (CCQs) are always incorporated in the questionnaires to evaluate attention towards processed videos based on specific factors (Dymarski et al., 2011). Inter-Observer Repeatability is defined next since it is the most popular reliability index; that is, assigning ratings to an identical item by different observers at different times. This will then show how much weight every ordinary user experience contributes to total output. Different transmissions or renditions lean towards an individual user experience—viewers' experiences—that normally ends up with a Total Quality Experience (TQE) score every time. Adaptation and qualities change over watching subsequently in viewing experiences.

Evaluation Management Sciences also shows the reliability exposure to the preservation of every subjected video-material i.e., dulcies, or specific elements like frame rates and codecs. Video-quality indexes before any re-coding are given for 2 hard-constrained codecs which extend to compressed length-volume, spatio-temporal-scales and density, dimensionality properties, and its own necessity. In addition aspect-ratio outcomes, delay-time periods as well as blurriness have a high impact on the final exams (Ivchenko et al., 2020).

4.3. Subjective-Objective Relation

The subjective-objective relation studies how different objective factors affect the subjective quality of experience (QoE) estimation in multimedia streaming. Despite the fast growth of objective metrics for multimedia streaming QoE, a subjective assessment is still necessary for properly evaluating the relation between the objective metrics and user-perceived QoE. The available standards cover the design of experiments on assessing quality through subjectivity; however, they do not provide detailed guidelines for conducting large validation experiments that correlate subjectivity with objective metrics, as usually required by multimedia quality assessment. Reception quality (QoR) models have proposed a unified performance model to represent the relationship between quality perceived by end users and delivered Qos indicator values of services being monitored (Ivchenko et al., 2020). Certain quality-of-service (QoS) metrics are highly redundant in measuring data packets and media streaming and have little impact on integrating new indicators into the QoR model for monitoring multimedia services and QoE-Recognizing Scheme. Quality-of-QoE models take a user-oriented perspective and offer real-time performance evaluation for supervising interactive and non-interactive service delivery to optimize the service better; hence they bear critical importance and have drawn increasing attention (Alreshoodi & Woods, 2013).

Analytically, objective–subjective modeling is typically addressed through (i) correlation analyses (Pearson, Spearman, or Kendall), (ii) mapping functions (logistic or nonlinear regression) that translate objective scores to MOS, and (iii) supervised learning models (e.g., random forests, gradient boosting, or neural networks) that integrate QoS, player events, and perceptual features. Reported correlations can vary considerably across datasets due to shifts in dominant impairments (such as compression versus stalling), which underlines the importance of evaluating models on multiple benchmarks and explicitly reporting the impairment distribution. A common gap in the literature is that numerous studies cite metrics without clarifying why they are effective (or ineffective) under specific streaming dynamics; consequently, subsequent sections highlight comparative strengths, limitations, and reproducibility requirements.

5. Experiment Validation and Benchmarks

Quality-of-Experience (QoE) metrics in multimedia streaming require experimental validation but with varying designs and many unspecified conditions. Two prominent studies assessing objective metrics for video-streaming QoE do not specify details about datasets, conditions, or results. An example of the advantages of dissemination is still pertinent to systematic planning. The methodology is described to show compliance with frameworks for experimental studies.

The commonly utilized public datasets and benchmarks in streaming QoE research and their respective capabilities, see Table 4. To enhance scientific rigor, studies should disclose (a) the dataset split strategy, (b) evaluation metrics (PLCC/SROCC/RMSE for prediction; QoE gain versus stall/switch trade-offs for control), (c) confidence intervals, and (d) reproducibility artifacts such as traces, player logs, and code. Cross-dataset evaluation is especially crucial, as models that perform well on a single benchmark frequently exhibit degradation when confronted with different content, devices, or latency regimes.

Table 4: public datasets and benchmarks in streaming QoE research

Dataset / benchmark	Typical scenario	Contains subjective scores?	Notable impairment factors	Where used / notes
ITU-T P.1203 open databases (feature-level)	HTTP adaptive streaming (VoD-like)	Yes	Stalling, quality switches, coding artifacts	Released with Robitza et al. (ACM MMSys 2018)
Waterloo SQoE-III	Adaptive streaming sessions reconstructed	Yes	ABR behavior, stalls/switches	Duanmu et al. (2018)
LIVE-Netflix QoE database	Mobile streaming-like conditions	Yes	Realistic network/buffer conditions	LIVE (UT Austin)
LIVE-NFLX-II	End-to-end streaming system with adaptation	Yes	Network and buffer conditions, adaptation effects	LIVE (UT Austin)

The consideration of QoE in streaming multimedia can be traced back to 2000 (Ivchenko et al., 2020). Visualiser platforms control substreams and specify content quality while mainstream providers emphasize delivery requiring decentralized control. Dynamic-adaptive streaming protocol development has changed distribution models and initiated research on video transmission under a goodput framework. Experimental campaigns proved that goodput is not significant in direct quality assessment though partial adherence to notation makes evaluation difficult; these limitations are present in existing databases for QoE.

Consolidation of datasets and provision of cross-laboratory benchmarks shall facilitate comparisons both within and across domains. Asset generators, which model alternative streaming protocols, remain outside the proposed framework. Detailed descriptions of setups promote reproducibility under different configurations. Proposals mention interest from parties involved in both adaptive-rate streaming and the subjective-objective correlation regarding trantransmitted video quality.

Further work includes statistical procedures for hypothesis testing and ranking, as well as the quantification of subjective-objective correlation over datasets and network configurations. Ongoing heuristic developments are to be validated empirically. Interest exists in adaptive-rate transmission plus related modeling paradigms; existing benchmarks plus the dataset identified support both research directions (Torres Vega et al., 2018).

6. Trends in QoE Research

The study of quality of experience (QoE) is a field that has been growing very fast with a lot of knowledge and applications. Much work has been done by the research community and industry to advance this area but the articulation of methods for practical implementation is still not enough. Some popular areas include adaptive streaming, immersive media, edge computing, and cross-layer optimization. Adaptive-streaming techniques come together with personalized content delivery in order to enhance user experience. Concerns about QoE with the quality of immersive media and its delivery characteristics have come up together with virtual reality (VR) and augmented reality (AR) devices becoming more popular. These are accompanied by the introduction of motion-to-audio-visual synchronization as a possible indicator for QoE, as well as the subjective assessment on whether VR/AR video is present or not. Edge computing plus distributed content-delivery networks have turned into mainstream solutions for latency reduction plus scalability improvement; however, they may also bring new challenges to QoE related to caching plus budget allocation for latency. Modeling QoE in multimedia streaming has shifted from experiment-to-model to cross-layer plus AI-driven approaches. There is an ongoing effort towards end-to-end optimization for QoE, as well as an attempt to bridge black-box machine-learning models with human-understandable factors of QoE (Alreshoodi & Woods, 2013).

6.1 Adaptive Streaming and Personalization

Adaptive streaming can be seen as a good answer for making use of resources while trying to bring out the best quality that a user could perceive. As bandwidth changes over time, adaptive algorithms will try to pick up the most appropriate streaming rate dynamically so that they improve interactivity. In one study about objective-quality measures taken over six video codecs, it was found that most measures had degraded performance during rate adaptation (Ivchenko et al., 2020). Another study found a strong correlation between start-up delay and subjective quality perceived by end-users (Tavakoli et al., 2014).

User preferences and contexts greatly affect the parameters of multimedia services that personalization strategies aim to adjust. Such preferences and contexts are the most important factors in determining the perceived Quality of Experience (QoE). The QoE of personalized services has not been extensively covered in literature, although evaluation is required for the design of effective service policies.

6.2. Immersive Media and VR/AR Considerations

The recent interaction with immersive media, augmented reality (AR), and virtual reality (VR) has huge effects on Quality of Experience (QoE) research. The media that is consumed through head-mounted displays, projection systems, or any other multisensory interface uses spatial audio, visual displays surrounding the viewer, haptic feedback as well as olfactory elements. This creates new opportunities for media streaming, gaming, and communication. New applications such as 360° video and volumetric video are responding to increased interest in immersive media. Starting with an overview of related requirements and user expectations, this section then reviews state-of-the-art approaches to motion-to-audio-visual synchronization and presence metrics for augmented and virtual reality (Dong & S. A. Lee, 2023).

The immersive experience provided by 360° virtual reality (VR) broadcasting of live music and sporting events captivates viewers but complicates content creation because multimedia streams become more complex than traditional 4K video. A spherical view requires transmission for 360° video resulting in high data demand and huge challenges for platforms like YouTube and Facebook. Quality-of-Service (QoS) metrics such as throughput and bandwidth have therefore received a lot of research attention. An increasing focus on Quality of Experience (QoE) acknowledges the human perspective and becomes more relevant in the context of emerging video technologies as well as the roll-out of fifth-generation mobile communication systems. Even so, growing interest in 360° content has not been reflected in scientific literature; hence subjective-assessment experiments need to be designed to create foundational knowledge (Covaci et al., 2019).

6.3. Edge Computing and Distributed Delivery

Quality of Experience is influenced by edge computing management; caching, content delivery network placement, and human-computer interactions are important for Quality of Service. Cache location as well as edge position should correspond to user demand while considering user mobility; therefore latency budgets should be relatively loose for real-life implementations (Alex Barakabitze et al., 2019).

Human-computer interactions also matter a lot when it comes to placement; distributed storage reduces initial access latency plus delays between pieces of content but must still fit into a predetermined latency budget if users want satisfaction with varying levels between videos.

6.4. Cross-Layer and AI-Driven QoE Models

Multimedia Streaming can be influenced by the available user connection capacity and device settings desired by the user. The different blocks in the video delivery chain, from video encoding to video decoding, provide different levels of service quality. It is therefore required to derive a set of Quality of Experience (QoE) models from the application point of view to support a scalable and reliable Endeavour streaming adaptive bit-rate video has a small number of layers. According to its definition, QoE gives priority to user expectation and perception of the streaming service. Macro motion is an important parameter both for fulfilling user needs and reducing modeling complexity, while motion in streaming services relates either to character movement or environment modulation (Ivchenko et al., 2020) .

7. Practical Implications for System Design

A comprehensive Quality-of-Experience (QoE) approach is required to systematically translate the understanding of user experience into practical engineering solutions and guide the design of next-generation multimedia streaming systems (Rodrigo dos Santos et al., 2023). The research-practice gap is growing more apparent with every passing day. As multimedia systems grow ever more complicated and as user expectations continue to evolve, it becomes critical that we enhance the interaction between objective QoE metrics on one side and subjective assessment on both sides in research as well as practice. Eight practical implications exist in closing this gap.

First, measurement infrastructure and tooling support are needed for real-time QoE monitoring and end-to-end experimental reproduction. Then, QoE-driven service-level objectives (SLOs), service-level agreements (SLAs), and policy specifications can help align network and application behaviors with user experience goals. Second, industry-standard evaluation protocols define consistent benchmarking requirements aligned with relevant regulatory frameworks; these protocols should facilitate comparative assessments of multi-dimensional solutions that optimize diverse cross-layer QoE objectives.

7.1 Infrastructure and Tools for Measurement

To ensure that the quality of experience (QoE) is continuously monitored and that experimental evaluations in streaming applications can be reproduced, a measurement infrastructure and tooling framework must be set up (Torres Vega et al., 2018). An infrastructure compliant with QoE must support the auditing of both objective and subjective metrics, integrate automation protocols, establish data-management buses and deploy end-to-end single-system measuring mechanisms (Rodrigo dos Santos et al., 2023). Comprehensive measurement systems are based on standardised formats for logging QoE-related data which are designed to allow replenishment according to pre-set triggers, enabling export onto different data management repositories for post-analysis. Instruments designed specifically for private experimentation and debugging give a push to research in fourth-fifth-sixth generation streaming services, testing newly developed metrics as well as comparative evaluation against alternative methods.

7.2 Service-Level Objectives from a QoE Perspective

Quality of Experience (QoE) is being increasingly adopted as the main metric for assessing perceived quality by users of multimedia streaming systems. Therefore, a number of objective metrics have been put forward to predict QoE while industries are trying to find ways to use these QoE-oriented metrics in their engineering practices. At the same time, new trends about QoE indicate good possibilities for future studies. Deriving service-level (SL) objectives from both objective and subjective metrics that will feed into designs and configurations over the whole life of a service brings great benefits. It takes Quality-of-Service (QoS), which does not really meet user needs, up to true quality-of-experience (QoE), which is one metric that can be guaranteed relevant to the user. One problem with establishing QoE-based SLs is that there is no cross-sector alignment on industry benchmarks or even on the basic goals of a service. Other problems — such as collecting measurements either continuously or systematically, knowing in real time the status of parameters needed to judge whether thresholds have been reached, and knowing how different metrics depend on each other — make things even harder.

QoE-centric SL objectives also enable policy exchange between organizations such as an ISP and a streaming service when negotiating rules for a streaming session. Such agreements usually reflect the fact that multimedia service stays within acceptable limits or just provides tolerance levels based on what has been designed into the offering.

7.3 Evaluation Protocols for Industry Standards

Evaluation protocols for industry standards are mostly concerned with Quality of Experience (QoE) evaluation in adaptive streaming over HTTP, which is a prominent delivery format for multimedia services. Objective factors such as bitrate, delay, distortion, and quality switching affect the subjective estimation of QoE; this is analyzed by Ivchenko et al. Regression models and correlation analysis are described, including the Gradient Boosting Regressor, adapted to applications with or without reference videos. The SQoE-III database used in this study overcomes some limitations present in earlier work, such as an inadequate number of samples and the absence of realistic conditions like buffering and quality switches. Traditional subjective assessment through expert ratings becomes impractical for large-scale video because of volume and bias.

8. Challenges and Open Questions

Quality of Experience (QoE) in multimedia streaming is a lively and ever-evolving research domain that seeks to offer user-centric quality assessment as a complement to traditional Quality of Service (QoS) evaluations. While QoS focuses on network and application performance levels, the objective of the QoE discipline is to assess how users perceive the quality of experience. Though extensively studied for streaming, QoE metrics will always be affected by a combination of client, server, network, and application parameters that ultimately determine user-perceived QoE. The wide gap between research efforts on QoE and its real-world deployment has triggered proposals for a new QoE Framework aimed at closing this gap and widening participation from interested communities in this collaborative effort.

9. Conclusion

Quality of Experience (QoE) for multimedia streaming encompasses a broad spectrum of user-centered factors, yet video/audio perceptual quality and assessment remain central to the scientific understanding and engineering of streaming playback quality, as several intertwined communities are generating progress framed in these terms. The focus on perceptual quality has grown since traditional QoS streaming parameters (image frame rate, audio sample rate, or number of service stalling interruptions) largely fail to correlate well with subjective user experience and the progressive nature of such applications.

QoE in multimedia streaming increasingly constitutes a multidisciplinary topic of critical importance that simultaneously supports the objectives of both academia and industry. Each party is concerned with various aspects of the user experience during and following the playback of streamed audio/video content, but the available solutions and characterizations differ significantly. Industry has consequently developed its own supporting set of models, metrics, and benchmarks that is in many respects distinct from that present in the current scientific literature.

References:

1. Alreshoodi, M., & Woods, J. (2013). Survey on QoE–QoS correlation models for multimedia services.

2. Barakabitze, A. A., Barman, N., Ahmad, A., Zadtootaghaj, S., Sun, L., Martini, M. G., & Atzori, L. (2020). QoE management of multimedia streaming services in future networks: A tutorial and survey. IEEE Communications Surveys & Tutorials, 22(1), 526–565.

3. Barman, N., & Martini, M. G. (2019). QoE modeling for HTTP adaptive video streaming: A survey and open challenges. IEEE Access, 7, 30831–30859.

4. Zhou, W., et al. (2022). A brief survey on adaptive video streaming quality of experience (QoE) assessment. Journal of Visual Communication and Image Representation, 103526.

5. Kougioumtzidis, G., Poulkov, V., & Zaharis, Z. D. (2022). A survey on multimedia services QoE assessment and machine learning-based prediction.

6. Robitza, W., Göring, S., Raake, A., et al. (2018). HTTP adaptive streaming QoE estimation with ITU-T Rec. P.1203 – Open databases and software. In Proceedings of the 9th ACM Multimedia Systems Conference (MMSys).

7. Duanmu, Z., Rehman, A., Wang, Z., & others. (2018). A Quality-of-Experience Database for Adaptive Video Streaming (Waterloo SQoE-III). IEEE Transactions on Circuits and Systems for Video Technology.

8. ITU-T. (2020). Recommendation ITU-T P.1203: Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport.

9. ITU-T. (2020). Recommendation ITU-T P.1204.3: Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to full bitstream information.

10. ITU-T. (2025). Recommendation ITU-T P.1204.2: Video quality assessment of streaming services over reliable transport for resolutions up to 4K using metadata and frame size information.

11. ITU-T. (2023). Recommendation ITU-T P.910: Subjective video quality assessment methods for multimedia applications.

12. ITU-R. (2023). Recommendation ITU-R BT.500-15: Methodologies for the subjective assessment of the quality of television images.

13. Netflix Technology Blog. (2016). Toward a practical perceptual video quality metric (VMAF).

14. Netflix Technology Blog. (2018). VMAF: The journey continues.

15. Netflix Research. (2020). Toward a better quality metric for the video community (VMAF).

16. Sheng, D., et al. (2024). PsyQoE: Improving quality-of-experience assessment with cognitive effects. IEEE Transactions on Services Computing, 17(6).

17. Zhu, Z., et al. (2024). Subjective and objective quality-of-experience evaluation for live video streaming. arXiv:2409.17596.

18. Du, L., et al. (2020). Video Quality of Experience metric for Dynamic Adaptive Streaming over HTTP (DASH). Applied Sciences, 10(5), 1793.

19. Tran, H. T. T., et al. (2021). Cumulative quality modeling for HTTP adaptive streaming. ACM Transactions on Multimedia Computing, Communications, and Applications.

20. Ghadiyaram, D., Pan, J., & Bovik, A. C. (2019). A subjective and objective study of stalling events in mobile streaming videos. IEEE Transactions on Circuits and Systems for Video Technology, 29(1), 183–197.

21. Covaci, A., Trestian, R., Comsa, I. S., & Ghinea, G. (2019). 360° mulsemedia: A way to improve subjective QoE in 360° videos.

22. Dong, H., & Lee, J. S. A. (2023). The metaverse from a multimedia communications perspective.